

Causal Modelling in Fertility Research: A Review of the Literature and an Application to a Parental Leave Policy Reform

Kreyenfeld, Michaela

Veröffentlichungsversion / Published Version
Zeitschriftenartikel / journal article

Empfohlene Zitierung / Suggested Citation:

Kreyenfeld, M. (2021). Causal Modelling in Fertility Research: A Review of the Literature and an Application to a Parental Leave Policy Reform. *Comparative Population Studies - Zeitschrift für Bevölkerungswissenschaft*, 46, 269-302. <https://doi.org/10.12765/CPoS-2021-10>

Nutzungsbedingungen:

Dieser Text wird unter einer CC BY-SA Lizenz (Namensnennung-Weitergabe unter gleichen Bedingungen) zur Verfügung gestellt. Nähere Auskünfte zu den CC-Lizenzen finden Sie hier: <https://creativecommons.org/licenses/by-sa/4.0/deed.de>

Terms of use:

This document is made available under a CC BY-SA Licence (Attribution-ShareAlike). For more Information see: <https://creativecommons.org/licenses/by-sa/4.0>

Causal Modelling in Fertility Research: A Review of the Literature and an Application to a Parental Leave Policy Reform*

Michaela Kreyenfeld

Abstract: This paper reviews empirical studies that have examined the causal determinants of fertility behaviour. In particular, we compare the approaches adopted in the different disciplines to improve our understanding of how birth dynamics are influenced by changes in female employment and changes in family policies. The wide array of panel data that have become available in recent years provide great potential for advanced causal modelling in this field. Event history modelling has been a dominant approach in sociology and demography. However, researchers are increasingly turning to other methods to unravel causal effects, such as fixed-effects modelling, the regression discontinuity approach, and statistical matching. We summarise selected studies, and discuss the advantages and the shortcomings of the different approaches. In an empirical section, we analyse the impact of the German 2007 policy reform on birth behaviour to illustrate the difficulties involved in isolating policy effects. The final chapter concludes by underscoring that even simple modelling strategies may be beneficial for improving our understanding of how policy effects shape demographic behaviour, and for laying the groundwork for more fine-grained causal investigations.

Keywords: Family policy · Event history · Fertility · Causation

1 Introduction

While most Western governments may shy away from formulating any clear-cut pro-natalist policy goals, they nevertheless have a vested interest in understanding the effects that family policies can have on birth dynamics. Scholars have approached this topic from different perspectives. Thus, the modelling strategies

* This article belongs to a special issue on "Identification of causal mechanisms in demographic research: The contribution of panel data".

which researchers use and the importance they attach to issues of causality have clearly been shaped by their respective disciplinary standpoints. The comparative welfare state approach has been among the most prominent concepts used to shed light on the interplay of family policies, female employment, and fertility behaviour – despite being the least committed to questions of causal modelling. Based on the simple assumption that women now “demand economic independence and permanent integration in working life”, it was argued that “familialistic regimes” would be punished by low national fertility rates (*Esping-Andersen* 1999: 69). *Esping-Andersen* has remained the most cited proponent of this line of thought, although similar ideas were developed contemporaneously by scholars from other disciplines (see e.g. *McDonald* 2000; *Rindfuss/Brewster* 1996).

The empirical evidence that *Esping-Andersen* (1999) generated in his seminal work relied on rather simple OLS regressions, whereby the outcome variable was the change in the national total fertility rate (TFR), while the independent variables were crude measures that depicted female employment rates and the degree to which social policies at the national level supported the compatibility of work and family life. Obviously, the strength of this concept did not emanate from its methodological rigour. Instead, it had filled a void that was left by other theories that failed to adequately address the pivotal role of family policies in influencing birth dynamics. Numerous studies followed that used more refined estimation techniques or employed more sophisticated measures that mapped the extent to which national policies help to achieve a work-family balance (*Ahn/Mira* 2002; *Castles* 2003; *Engelhardt/Prskawetz* 2004; *Luci-Greulich/Thévenon* 2013; *Thévenon/Gauthier* 2011).

Overall, the evidence from macro-level data remained inconclusive and unsatisfactory (*Gauthier* 2007). As the unit of analysis is the country, only a few observations could be included in such investigations, which resulted in unstable results that relied on a set of rather heterogeneous countries. Most importantly, concerns were raised about the outcome variables, which were often restricted to simple summary indicators, such as the national total fertility rate (TFR). The conventional TFR conflates different birth parities. But even if vital statistics contain a parity-specific TFR, these data are still limited as they only rarely include real “exposure data” that relate births to the population at risk of a certain birth parity. Furthermore, these types of data do not allow researchers to conduct fine-grained analysis; for example, to examine how different subpopulations react to a policy change. Longitudinal micro-level analyses were deemed to be a more promising avenue in this area of research as they were less prone to the well-known problems associated with ecological fallacies, and they opened up new possibilities for unravelling the causal effects behind birth behaviour.

This paper discusses the contribution of longitudinal micro-level analysis to the understanding of fertility dynamics. We provide an overview of two broad streams of literature in this area. First, we discuss studies that have examined the causal impact of women’s work on fertility behaviour. Second, we draw attention to studies that have explored how national policies have influenced birth dynamics. The third part of the paper takes the German 2007 policy reform as an example to illustrate the

difficulties researchers face when seeking to isolate policy effects in an empirical investigation. The final section summarises and evaluates the pros and cons of the different approaches. Such an overview can never be complete. It is limited to the two abovementioned broad research streams, and is restricted to studies that use the potential of longitudinal micro-level data to unravel causal effects, which may be panel surveys, retrospective surveys, or process-produced administrative data.¹ Furthermore, most of the literature we discuss below addresses behaviour in European countries since the 1990s.

2 Female employment and fertility

2.1 Contributions and limitations of event history modelling

Despite claims that the heyday of regression analysis has long since passed (*Morgan/Winship* 2015), event history modelling has remained a dominant method for studying fertility behaviour, at least in demographic and sociological research. Unlike in simple cross-sectional OLS regression, “time” is at the heart of these methods. Through its use of *time-varying covariates*, event history analysis lives up to one of the main principles of causal analysis, namely that the cause should always precede the effect (*Bhrolcháin/Dyson* 2007). Thus, a large body of literature emerged that examined how women’s work – operationalised by a time-varying covariate distinguishing between different employment states – is related to first-order and higher-order birth risks (for a meta-analysis, see *Alderotti et al.* 2021; *Matysiak/Vignoli* 2008). Because the moment of childbirth is not equivalent to the decision to become a parent, these studies conventionally predated the date of childbirth by several months. Important policy-relevant results were generated from this type of research, which illustrated that the links between female employment and fertility can vary greatly across policy contexts, socio-economic statuses, birth parities, cohorts, and time periods (*Andersson et al.* 2014; *Matysiak/Vignoli* 2008, 2013; *Schröder/Brüderl* 2008; *Wood/Neels* 2017).

Micro-level investigations also opened up the opportunity to operationalise employment status in a more sophisticated manner. While most of the macro-level studies were confined to simple summary indicators, such as the female employment rate, micro-level studies could operationalise women’s work over a fine-grained measure that differentiated between various employment categories, including educational participation, gainful employment, unemployment, and other activities. While this level of granularity enabled researchers to cast a more nuanced

¹ This overview is limited to studies that employed longitudinal data. However, we must pay tribute to the earlier empirical studies that tried to unravel causal effects, even though they relied mostly on cross-sectional datasets only (*Cramer* 1980; *Stolzenberg/Waite* 1977; *Thomson* 1984; *Waite/Stolzenberg* 1976). In addition to laying the groundwork for conceptualising the association between female employment and fertility, these studies motivated much of the later collection of longitudinal data.

light on the micro-level correlations between female employment and fertility, it also raised broader questions about the meaning of women's employment status. This was particularly evident for the category of "housewife" that was included in some of the datasets. It was deemed obvious that housewives would experience elevated first birth rates, based on the assumption that their life course concept was closely intertwined with the idea of having a family. Although the time-varying modelling of women's employment status guaranteed a correct ordering of cause and effect, it was also evident that it did not entirely eliminate *reverse causality*. People plan their lives ahead and they base their current actions on their anticipated future behaviour (Hoem/Kreyenfeld 2006). An individual's current employment status is therefore more than just an indicator of his/her present circumstances. Instead, it is the product of choices that factor in anticipated future behaviour and thus the individual's "personal imaginaries" or "narratives of the future" (Bazzani et al. 2020; Vignoli et al. 2020a/b).

Another source of bias in event history modelling comes from the modelling of employment as a time-varying covariate. In such cases, *endogenous selection bias* may bias the coefficient for past employment status, if past employment affects the confounders of current employment (Elwert 2013; Elwert/Winship 2014; Sharkey/Elwert 2011). Multi-process models or simultaneous equation event history models have been proposed in an effort to address endogeneity (Lillard 1993). The advances in such methods have been closely connected to the development of the software package aML (Lillard/Panis 2000), which led to an upsurge in the use of these methods in family research, particularly in the early and mid-2000s (e.g. Aassve et al. 2006; Upchurch et al. 2002; Kravdal 2001; Kulu/Steele 2013; Schnor 2014). This modelling approach is based on the premise that partnership formation, marriage, educational attainment, and fertility behaviour are transitions that may influence each other as these processes unfold. Multi-process modelling tries to capture this endogeneity by letting the outcome of one process (such as the formation of a partnership) influence the other process (such as educational attainment), and by including heterogeneity components in each process. These heterogeneity components are usually assumed to follow a joint normal distribution. A methodological limitation of these models is that they hinge on this underlying distributional assumption.² While this assumption may be relaxed, employing alternative ones often leads to practical problems, such as the models failing to converge (see e.g. Kravdal 2001: 198). A more substantive limitation of these models is that the unobserved heterogeneity components remain a black box, with researchers being unable to pinpoint the exact mechanisms that lead to a particular behaviour. Furthermore, even if employment and fertility can be modelled as mutually related processes, such efforts will not eliminate the bias that arises because people are acting on their anticipated behaviour in the future (the abovementioned "personal imaginaries").

² Additionally, the model assumes that the source of unobserved heterogeneity is constant across time, which may also be a strong assumption.

Endogenous selection bias is conceptually different from the more widely known problem of *common-cause confounding*, whereby the bias comes from having failed to include a variable that causes both the variable of interest (employment) and the outcome variable (fertility) (Elwert 2013; Elwert/Winship 2014; Hernán et al. 2002). A typical example would be the omission of migration background, which may affect both fertility behaviour and employment. Standard regression analysis usually tries to combat this problem by including additional variables in the model. Thus, the main strategy for assessing causality in regression models is to eliminate “alternative effects” by controlling for confounders. This strategy has several implications. It hinges on the availability of variables, whereas some variables may simply not be available in a given dataset. In addition, the results may vary depending on the number and the type of variables that are included in the regression. This is also one of the main conclusions reached by Matysiak and Vignoli (2008) in their meta-analysis of 30 journal articles that used event history modelling to assess the effects of female employment on fertility. The authors concluded that the prior results were often unstable and heavily influenced by the inclusion of available covariates. While this is undoubtedly a valid conclusion to draw from this meta-analysis, it can leave researchers in some despair. Scholars may be torn between including further covariates and ending up with a grossly over-specified model, or choosing a parsimonious modelling strategy while running the risk of being accused that the model does not sufficiently account for all possible confounders.

2.2 Fixed-effects modelling for fertility research?

The self-selection into “treatment” is one of the main reasons why observational studies will always remain limited in their ambitions to make causal claims. While experimental designs, such as randomised controlled trials, are usually considered the gold standard for causal analysis, conducting them remains an elusive goal for most demographic and sociological researchers, as ethical reasons generally prohibit the use of such designs (Bhrolcháin/Dyson 2007). Nonetheless, most of the progress made in recent decades in the development of sociological research methods has been geared towards mimicking an experimental design approach (Gangl 2010). The terminology of a “control group” and a “treatment group” clearly signals such a commitment. The rise of the use of fixed-effects regression modelling has been informed by this broader aim (Brüderl/Ludwig 2015; Hill et al. 2020). Fixed-effects modelling is based on the idea that a person can be used as his/her “own control group” if longitudinal data are available (Allison 2009; Brüderl/Ludwig 2015). Thus, fixed-effects panel regressions make causal claims by exploiting the variations that exist within a person over time. A classical demographic example that demonstrates the power of this method is the investigation of the relationship between marriage and life satisfaction (Brüderl/Ludwig 2015). In contrast to a standard regression approach that compares the life satisfaction of married and unmarried people, this method follows a person throughout his/her life course, and assesses how the individual’s life satisfaction changes when s/he experiences the

event of marriage. In this context, marriage is considered the “treatment”, while the “control group” is the individual before s/he experienced the event. Technically, this method eliminates all time-constant heterogeneity that commonly distorts standard regression analysis (*Brüderl/Ludwig* 2015).

Many studies have employed fixed-effects panel regression to study different demographic and sociological themes, such as the demographic determinants of life satisfaction and other dimensions of well-being. Moreover, childbirth has been used regularly as an independent variable in these models (e.g. *Myrskylä/Margolis* 2014). It is noticeable, however, that only a few studies have employed this method to examine childbearing behaviour as an outcome variable. In the few cases in which the method was employed for fertility research, the results were mainly published in economic journals (*Boca et al.* 2005; *Huttunen/Kellokumpu* 2016; *Michaud/Tatsiramos* 2011).³ Have demographers and sociologists been too slow to catch up with this methodological innovation? Or are there sound reasons why this method has not yet been widely adopted in fertility research?

A possible answer to these questions is that, the study of fertility from a parity-specific point of view is almost an imperative in demographic and sociological research. The decision to have a first child is regarded as a distinct choice that is determined by factors that are different to those that result in higher-order births. If we subscribe to this idea, we need to model the transition to first parenthood as a unique event. Fixed-effects panel regression runs into a problem in such cases, as it is based on the premise that the event (i.e. the birth) should be observed before and after a treatment (e.g. a change of employment status). If an event is observed only once in a lifetime, there is a fundamental flaw in the logic as there is no before and after. If the event of the first birth occurred before the “treatment”, it cannot be repeated afterwards.

To get around this problem, some researchers have treated the number of children as a continuous variable (see e.g. *Boca et al.* 2005; *Huttunen/Kellokumpu* 2016; *Michaud/Tatsiramos* 2011). Obviously, this approach conflates different birth parities. But there are other concerns as well. Births are infrequent events in contemporary European societies, and people rarely have more than two or three children. As such, there is little variation in the outcome variable, and some of the variation comes from the more “unusual” cases of people with larger families. Moreover, employment and fertility are events that strongly influence each other. The birth of a first child often results in drastic changes in a woman’s work pattern. A large share of women do not work shortly after giving birth, but re-enter the labour market at some later point in time. In other instances, women may space their first and second children closely together to minimise their family-related employment interruptions. The complex interplay that exists between fertility and women’s employment gets completely lost in these models, such that it ultimately becomes

³ This observation pertains to panel fixed-effects modelling. Fixed-effects modelling has been regularly used with macro-level fertility data in sociological and demographic fertility research; for example, to account for regional-specific effects (e.g. *Schneider* 2015).

difficult to grasp the meaning of the underlying behavioural patterns. Thus, the use of fixed-effects panel regression is not a panacea for causal analysis (*Collischon/Eberl 2020; Hill et al. 2020*) and it is clearly not the best way to move forward in the analysis of fertility behaviour.⁴

2.3 Fertility intentions and panel data

While fixed-effects modelling may not be the preferred method for examining *fertility behaviour*, it can be applied more effectively to the analysis of *fertility preferences*. The opportunities to analyse fertility preferences with European panel data accelerated in the early 2000s, when data became available from sources such as the Generations and Gender Survey, the British Understanding Society study, and the German Family Panel, which included recurrent items on the respondents' fertility desires, ideals, and intentions.⁵ *Kuhnt, Kreyenfeld, and Trappe (2017)* used data from the German family panel (pairfam) to study changes in the *ideal number* of children across the life course by means of fixed-effects modelling. The authors did not find any association between women's employment status and their ideal number of children. The only variable that was shown to have a substantial effect was "having (further) children", which suggests that the respondents were adjusting their ideals to align them with their behaviour. *Kuhnt, Minkus, and Buhr (2020)* employed the same data, but used the intention to have a child (within the next two years) as an outcome variable in a multinomial fixed-effects regression model (with the categories "certainly yes", "uncertain", and "certainly no"). While the authors did not find that a woman's employment status affected her fertility intentions, the model showed that men's worries about finding a suitable job increased the likelihood of being uncertain about the plans for having a child within the next two years. A non-trivial problem that tends to arise in this type of investigation is that some respondents may realise their fertility intentions and have a child during the study period. As such, there is selective drop out from the study population.⁶

Other studies have focused more specifically on the link between intentions and fertility behaviour. *Berrington and Pattaro (2014)* used data from the UK data

⁴ *Schröder and Brüderl (2008)* used an event history model to study first birth rates. Instead of controlling for employment status, they controlled for whether a person had changed his/her employment status (i.e. moved from non-employment to employment). In doing so, the authors used standard event history analysis, but also exploited the within-variation.

⁵ The collection of large-scale longitudinal data that included items that measured fertility desires, intentions, and ideals was initiated in the US as far back as the 1970s. Much of the work on fertility intentions was motivated by studies based on US data (see e.g. *Hayford 2009; Thomson 1997*).

⁶ A downturn in these investigations is that all birth parities were pooled in this investigation. Thus, the fertility intentions of a particular woman that is observed across time may refer to different birth parities. Furthermore, problems of selectivity may arise, because women who did not state a preference, or who were pregnant at the time of the interview, are excluded from the sample. In a similar vein, the realisation of an intention (childbirth) is not adequately accounted for in this model.

“Understanding Society” to compare women’s fertility intentions at age 23 with their actual number of children at age 46. This study found that highly educated women were less likely to realise their intentions. However, the results also showed that a woman’s employment status at age 23 did not seem to be related to her chances of realising her early fertility intentions. Other studies examined how a woman’s fertility intentions as measured in a given year determined the probability of the woman realising those intentions by the next year or the next round of data collection (e.g. *Hanappi et al.* 2017; *Kuhnt/Trappe* 2016; *Riederer et al.* 2019; *Spéder/Kapitány* 2009). These studies have greatly advanced our understanding of the obstacles women face in realising their fertility plans. Ultimately, however, these models do not solve the problem that a woman may self-select into a certain employment status across her life course.

2.4 Natural experiments

Our interest in understanding the causal effects of female employment on fertility stems from the fact that in economic and demographic theories of fertility, women’s labour market participation has been seen as one of the prime factors contributing to declining or low fertility rates. Refuting the hypothesis that women’s employment leads to lower fertility not only challenges the predictions from conventional theories; indeed, in the narrow economic sense, it means that the “income effect” has ultimately trumped the “substitution effect”. More broadly, it implies that women’s role in society, the meaning of female work, and the economic foundations of the family have all shifted. As these are strong implications, it is imperative that we rule out the possibility that selection into employment has produced the observed outcomes.

A standard approach used in economics to establish causality is to search for *natural experiments*; i.e. events that affect the variable of interest (female employment, non-employment, or unemployment), but that are beyond the control of the individual. Examples of such natural experiments are large-scale lay-offs, bankruptcies, and firm closures. These events have been used previously in labour market research to examine how unemployment, operationalised over these exogenous unemployment shocks, alters people’s behaviour and well-being. However, this approach has seldom been used in fertility research (see, however, *Del Bono et al.* 2015; *Hofmann et al.* 2017; *Huttunen/Kellokumpu* 2016).

With regard to the small number of existing studies that adopted this approach, most were conducted by labour economists who employed *matching techniques*. The underlying idea of this approach is that true causal analysis requires a comparison of a control group and a treatment group. Standard observational studies suffer from the problem that the comparison group may be select, and is thus not comparable with the treatment group. To ensure comparability, matching techniques start with the sample of “treated” individuals and search the pool of the control group for those units that best correspond to the treatment group. This “search procedure” has been refined in recent years by employing sophisticated algorithms, most recently from machine learning operations (*Lee et al.* 2010). However, despite recent

refinements in these approaches, researchers using matching techniques struggle with problems similar to those faced by researchers using conventional regression models. To ensure a good match, they need a sufficient number of covariates. If these covariates are not included, the matching techniques will result in poor matches – and, by extension, in biased results.

In addition, matching techniques are not always easy to apply to demographic processes. *Hofmann and colleagues* (2017) examined, for example, the causal effects of unemployment on first birth risks. The paper compared the subsequent birth risks of individuals who lost their jobs after having been laid off by a firm with those of a control group of women who were not subject to a firm closing. To make the control and the treatment groups comparable, firm-level information was required. As a result, the sample was limited to respondents who had been employed with the same firm for at least 1.5 years. Women who had been working at firms with more than 2,000 employees or at very small firms had to be excluded as well, because it would have been difficult to find suitable matches for these women. In the end, after the restrictions were imposed, the original sample of 101,910 individuals shrank to 3,286 treated and 17,836 untreated women. Although the sample restrictions were clearly spelled out and clear justification given for them, the results from such an analysis are inevitably limited to a restricted subpopulation. In this case, one of the restrictions required by the sample was that the women were employed with the same firm for at least 1.5 years. This limited the results to women who were firmly established in the labour market. Thus, while this model may have generated careful causal effects, the external validity of the results was limited.

2.5 Narratives of the future

A new impetus for understanding the causal link of the “female employment and fertility nexus” comes from studies that have tried to integrate future uncertainties into models of rational decision-making. These models break from the presumption that many decisions, including the decision to have children, are long-term commitments. It is therefore argued that actors need to take into account the uncertainties regarding the future implications of their choices. Some future states of the world may be extrapolated from the “shadow of the past” (i.e. it may be assumed that a woman’s husband will continue to have a high income if he is currently earning well). Other states of the world are more difficult to predict, and actors may “play” with imaginaries to deal with these uncertainties (*Beckert/Bronk* 2018).

“Narratives” are standard concepts in psychology and sociology, and have recently also been employed in economics (*Bénabou et al.* 2018; *Shiller* 2020) and demographic research (*Vignoli et al.* 2020a). It has been argued that while conventional fertility models have primarily dealt with the effects of present states, fertility choices are governed by narratives of the future, which stem from the “human capacity to place oneself in an imagined situation that cannot be deduced from present conditions” (*Vignoli et al.* 2020b). In other words, the decision to have

a child depends on how a woman visualises the future; for example, whether she can imagine being a “working mother”, or whether she can only imagine being a “homemaker”. This area of research is still evolving, and among the major challenges scholars have faced is that of generating suitable item sets that can capture these “narratives of the future” (Vignoli *et al.* 2020b) or “images of the future self” (Bachrach/Morgan 2016: 466). What is interesting about this debate is that it shifts attention away from developing sophisticated econometric techniques for studying the causal effect of employment on fertility, and instead focuses on finding appropriate item sets to measure future “models” or “narratives” of the world. It has also been proposed that more qualitative research should be conducted, including longitudinal qualitative research that helps to elucidate how these “narratives” are formed and altered across time (Bernardi 2021; Vignoli *et al.* 2020b).

The concept of “future narratives” resonates well with the considerations developed by Friedman, Hechter, and Kanazawa (1994). They provided a rationale for why some women may decide to have children in a seemingly unstable and insecure economic situation, and why this decision may still be regarded as rational. On the one hand, childbirth tends to block biographical “alternatives”. Thus, having children while unemployed may limit a woman’s chances of swiftly returning to the labour market and finding a job with good career prospects. On the other hand, this limitation may be perceived as a relief and a source of “uncertainty reduction”, as it structures a woman’s otherwise uncertain life course.⁷ For women who foresee that they will not be able to succeed in the labour market, taking on the role of a homemaker may be regarded as a meaningful and socially accepted “biographical alternative” to an uncertain labour market trajectory. Clearly, the societal context matters here because it defines the available “narratives” or “biographical alternative”. For example, policies may encourage or discourage a certain earner model; or the normative fabric of a society may define gender roles, attitudes towards non-parental care, and, more generally, attitudes towards parents’ roles as earners and carers.

3 Policy change and fertility

3.1 Isolating policy effects

In response to declining and low birth rates in European countries, policy-makers have become increasingly interested in the question of whether policies could halt the downward trend in period fertility and reduce childlessness, particularly among highly educated and work-oriented women. In this debate, it was argued that the expansion of publicly financed child care and the implementation of earnings-related parental leave benefits are key measures that can simultaneously support

⁷ Lehrer *et al.* (1996: 133) criticised this argument for the assumption that the risks “emanating from parenthood are more controllable than those associated with labor force participation”.

higher levels of gender equality and fertility. Against this background, the European Commission issued recommendations to expand public day care and to implement suitable parental leave regulations (Annesley 2007). Although many national governments followed their own agenda and did not necessarily comply with the EU recommendations, most European governments have scaled up their public day care services and implemented parental leave regulations since the 1990s (Daly/Ferragina 2018). Child care services have been gradually expanded, and the speed of this expansion has depended in part on the capacities and conditions at local municipality level. Researchers have thus been able to use the regional variations in these child care expansion efforts to isolate the possible policy effects (e.g. Bauernschuster *et al.* 2016; Hank/Kreyenfeld 2003; Krapf 2014; Kravdal 1996).⁸ However, as parental leave benefit regulations were generally implemented at the national level, scholars had to use other strategies to examine the impact of these measures on birth behaviour.

Much of the early demographic research into the causal impact of policy changes on fertility behaviour has dealt with developing economies, and thus with high-fertility settings. For these regions, *randomised controlled trials* have been regularly employed to evaluate the effectiveness of family planning programmes (Mwaikambo *et al.* 2011). For advanced economies, these methods are applied less commonly in the area of fertility behaviour. There have, however, been investigations of how sexual education (and other “moderate” interventions) have affected teenage pregnancy rates (for an overview, see Bennett/Assefi 2005). Ethical issues, including concerns about equal access to public services, often prevent the implementation of randomised control trials to evaluate the impact of more coercive policies on fertility behaviour (Bhrolcháin/Dyson 2007).

However, the programme evaluation literature offers a battery of techniques for isolating policy effects with observational data (LaLonde 1986). A standard method that is often used is the *difference-in-difference (DiD)* approach, which is commonly combined with the abovementioned matching techniques. This approach suffers from the same drawbacks as fixed-effects modelling (see Section 2). The DiD requires an observation before and after the intervention to isolate the causal effect of a policy. As the first childbirth is an absorbing event, DiD cannot be employed to study first birth processes. Researchers could combine all birth orders so that they would be able to observe events before and after treatment (a policy reform). However, this would violate a basic understanding of demographic research: namely that first and higher-order births are distinct processes.

In light of these challenges, researchers have increasingly turned to the *regression discontinuity approach* to examine how policy reforms impact behaviour. The regression discontinuity approach is based on the idea that causal inference can

⁸ The assumption here is that regional and temporal differences in child care slots depend on the capacities and willingness of local actors to increase child care services. If municipalities offer child care services based on expected demand, regional changes in provision rates cannot be used as a suitable exogenous policy measure, as they already reflect the usage of care services.

be generated by comparing the behaviour of individuals immediately before and after a cut-off date of a policy reform. This method has, for example, been used to examine how changes in parental leave regulations have affected labour supply (Ginja *et al.* 2020), living arrangements (Cygan-Rehm *et al.* 2018), as well as childbirth (e.g. Cygan-Rehm 2016; Farré/González 2019; Tamm 2013). While these models are central to the field of policy evaluation, they have high data demands. In order to generate significant results, large sample sizes that include reasonable numbers of events that are observed immediately before and after a policy reform are needed. While this need for large amounts of data is a general concern when using these types of models, it is a particularly pressing problem in fertility research as births are rare events. Therefore, most survey data will not provide a suitable basis for applying the regression discontinuity approach to birth behaviour. Consequently, this method can mainly be used in fertility research if large-scale register data are available.

3.2 Criteria and considerations for the analysis of policy effects

In order to study policy effects, suitable *counterfactuals* are needed. The Nordic countries have been ideal for the counterfactual approach. These countries are at the vanguard of progressive parental leave regulations, having already overhauled their family policies in the 1970s and 1980s. Furthermore, as these countries have large-scale register data suitable for individual-level fertility analysis, studies conducted in these countries were among the first to generate solid evidence of how parental leave policies could alter fertility patterns (Andersson 1999; Andersson *et al.* 2006; Björklund 2006; Neyer/Andersson 2008). A stylised finding from this body of research is that the introduction of the so-called “speed-premium”⁹ in the Swedish parental leave system led to an increase in second birth rates, and a shortening of birth intervals. This evidence was generated based on simple parity-specific birth rates (i.e. from event history models) for Sweden that were compared with patterns in other Nordic countries. In the Nordic context, where there are only minor differences in the cultural and economic developments between countries, it seems straightforward to analyse Swedish fertility and to use neighbouring Norway or Denmark as counterfactuals. Other European countries often lack such clear-cut comparison groups, because the data are often less comparable across countries; and because countries differ on multiple dimensions, including in their cultural and economic conditions, and in their fertility patterns prior to a reform.

In addition to the challenge of guaranteeing similar prior conditions, Bhrolcháin and Dyson (2007) have listed further criteria for performing meaningful causal analysis of demographic outcomes. In particular, they have emphasised the importance of

⁹ Parental leave benefits in Sweden cover roughly 80 percent of prior income. If the first and second births are spaced closely apart, women will receive the same parental leave benefit for the second child as for the first one, even if they had reduced their employment and earnings between the two births. The same applies to higher-order births (Andersson *et al.* 2006).

contiguity: i.e. that the effect should follow shortly after the cause (*Bhrolcháin/Dyson* 2007: 8). Moreover, the regression discontinuity approach (see above) capitalises on the idea that a policy reform can be regarded as a sudden external shock that leads to an immediate reaction. Sharp cut-off dates for eligibility for certain benefits are very important tools in understanding how policies operate. However, not all policy reforms are implemented in a clear-cut fashion. Furthermore, there may be a mismatch between the *de jure* regulations and the *de facto* implementation of a policy. Some groups may be eligible for a service or payment, but may not be able to take advantage of it, either because they are simply not aware of their eligibility or because they are unable to master the bureaucratic hurdles involved in applying for a state benefit. Moreover, in family law, there is often a “margin of appreciation” of how legal texts are interpreted and enforced in practice. This is most evident in the regulation of child custody or alimony, in which judges have a certain leeway in their rulings. There is only a small margin of appreciation in the case of parental leave regulations, as eligibility is clearly defined in law. However, for some families, bureaucratic hurdles may have indeed been a barrier to taking advantage of these programmes. The abovementioned parental leave benefit reform requires parents to document their earnings of the last 12 months. The German Family Ministry has only recently intensified efforts to make the procedure easier (by retrieving the earnings information automatically from the pension registers). Returning to the abovementioned principle of *contiguity*, we observe that this principle seems to conflict at least somewhat with the ways in which some policies play out in practice. Thus, there may not be an immediate “reform effect” because it can take time for the knowledge of how to apply for certain benefits to trickle down to all layers of society. In addition, governments may need time not only to optimise their systems and thus remove bureaucratic hurdles, but also to make the application procedure easier so that access is guaranteed even for hard-to-reach populations, such as those who do not speak the native language.

Also relevant in this context is the *policy process*. From the ministerial draft, to the discussion in parliament, to the enactment of the law, and to its ultimate implementation; this process can easily stretch across several months, or even years. Depending on the magnitude and relevance of the reform, this debate may receive media coverage. As a result, even if we are able to pin-point the exact date when a policy reform came into force, the reform may have increased levels of public awareness much earlier. For example, Germany drastically curbed ex-spousal maintenance after divorce in 2008 (*Radenacker* 2020). As this policy process was covered in the media, it is likely that women who foresaw that they would get divorced had altered their employment behaviour well in advance of the actual date of the implementation of that reform. The parental leave benefit reform, which will be analysed below, came into force on 1 January 2007. The reform bill was passed on 20 June 2006.¹⁰ It is possible that some couples foresaw that this policy would

¹⁰ <http://dip21.bundestag.de/dip21/btd/16/018/1601889.pdf>

come into effect, and acted accordingly. Thus, births in January 2007 may have been planned in anticipation of the implementation of the reform. Therefore, an immediate reaction to the reform may have been observed. However, given that births are not that easy to time for most couples, a more realistic scenario is that there was a certain time lag between the implementation of this reform and its effects on fertility.

Policies are rarely formulated in a vacuum, and how they operate also depends on the context. In some cases, a policy may not succeed because people do not use it. Levels of *policy usage* may vary for different reasons. For example, eligible individuals may not consider applying for a transfer for fear of being stigmatised if they collected the benefit. Similarly, parents may not take advantage of an available child care slot because societal norms sanction the use of day care at young ages. *Neyer and Andersson* (2008: 702) argued that certain policies may be deemed to have failed as a result of a lack of *coherence* and a mismatch with the broader societal system. For example, the German parental leave benefit reform was basically copied from the Swedish system. Unlike Sweden, which had already abolished individual taxation in the 1970s, Germany retained its system of joint taxation. The German system continues to provide tax relief if the “second earner” (usually the woman) has worked less. Thus, the system does not provide a clear commitment to the “dual earner model”, but instead provides strong incentives to opt for the single earner model. *Kalwij* (2010: 704) concluded that researchers should never analyse policies in isolation, nor should they “simply sum up the various family policies”. Instead, researchers should consider the interplay of the different measures and how they align and resonate with the overall logic of the system.

Neyer and Andersson (2008) introduced the concept of *critical junctures* into this debate. This terminology was originally used in the political economy literature to depict turning points in institutional developments. Critical junctures are policy reforms that have a lasting effect on the entire welfare state framework because they determine its future pathway and can “lock” the system into moving in a certain direction (*Pierson 2000; Thelen 1999*). It also follows that the constellation of different policies generates a unique and separate dimension. In addition to pointing to this idiosyncratic nature of policies, the authors suggested to analyse possible interaction effects between the existing policy measures and the newly introduced ones. For example, it is conceivable that parental leave policies would only be effective if they were combined with an increase in the provision of child care. While child care policies could be integrated by exploiting regional variations, other policies are more difficult to integrate (such as the effects that come from the tax and transfer system). The use of large-scale cross-national data, combined with information on the features of the different welfare states, could provide some important insights into how policies operate in different contexts (*Wesolowski et al. 2020*).

4 How did the German parental leave benefit reform affect first birth behaviour?

4.1 The 2007 parental leave benefit reform

Below, we use the German parental leave benefit reform of 2007 to illustrate the difficulties involved in isolating policy effects. The German parental leave benefit reform came into force on 1 January 2007. The reform, which was copied from the parental leave benefit programmes of the Swedish system, was regarded in the academic community as marking a major departure from the previous system in Germany (*Fleckenstein* 2011; *Spieß/Wrohlich* 2008). The prior regulations provided a flat-rate benefit of 300 euros per month for the duration of two years, while the new regulations provided a shorter term of parental leave of only 12 months. Moreover, the benefit was now earnings-related, with the net earnings in the year prior to childbirth serving as the basis for the calculation of the benefit.¹¹ Although the term during which benefits were received was shortened, the new regulations generally led to an increase in monthly payments, one exception being the unemployed and persons who were not integrated into the labour market. The benefits of these individuals were often lower under the new system.

Some scholars characterised the reform as a major and radical shift and a clear departure from the logic of the previous system, which was often classified as a conservative and familialistic welfare state regime (*Fleckenstein* 2011). In addition, the new regulations provided strong incentives for women to become established in the labour market before having children. Thus, a clear-cut hypothesis follows from these observations: namely, that the association between female employment status and earnings should have changed after the reform, with women becoming more likely to postpone childbearing until they were firmly established in the labour market. This hypothesis will be tested below based on large-scale register data.

4.2 Data and variables

Data for this investigation come from the German Pension Registers. We use the “VSKT 2015”, which is a sample drawn from the registers and which includes persons with German citizenship who had an active pension account in 2015, i.e. persons who were not yet retired (*Stegmann* 2018). The main benefits of using this dataset are its large sample size and its detailed and reliable employment, earnings, and fertility biographies (for a validation of the fertility histories, see *Kreyenfeld/Mika* 2006). There are, however, several disadvantages to using this dataset. First, the pension data cover only about 90 percent of the resident population in Germany, as opposed to the total population. People in certain professions (farmers, civil servants) are not included in the registers. An additional drawback is that it contains

¹¹ There is an earnings cap for the calculation of the benefit (currently 1,800 euros).

only a few variables that can be employed for the investigation. For example, partner information or information at the household level are not available in the dataset. Thus, the analyses must rely on a “sparse model” with a limited number of covariates. Another drawback is that the data only include individuals born between 1948 and 1985. This means that coverage of younger ages for the very recent time period is poor.

The data have been transferred into a person-month dataset, and have been restricted to women aged between 20 and 40 in the years 2005 to 2010. We do not analyse earlier years because unemployment is one of our main covariates, and because measures of unemployment are not fully comparable for longer time periods due to a major policy reform of the unemployment benefit scheme that came into force in 2004 (the so-called Hartz IV reforms). We do not analyse later years because coverage of very young women for the period after 2010 is poor. Sensitivity analyses were conducted that cover longer time periods for older women (see Fig. A2 in the appendix).

The main outcome variable is the *first childbirth*, which was pre-dated by nine months. Women who had a child before the observation window are not part of this investigation. Cases are censored at “last clearance”, which is the date when the German pension fund contacts a person to verify the information provided in the registers. The total number of women in the analytical sample is 48,843, which corresponds to 2,833,078 person-months. These women gave birth to 13,913 first children in the observation window (see Table A1 in the appendix for the sample statistics).

The main covariate is the *employment status*. Employment status is available on a monthly basis in the registers. We distinguish between women who were in (1) schooling, (2) employment, (3) unemployment, and (4) others. *Schooling* includes any educational episode, including school attendance, vocational training, and university education. Not all educational episodes are equally relevant for pension payments. For this reason, university education (as well as further education) is not fully recorded in the registers.¹² Thus, some educational episodes cannot be identified, and they will appear in the “other” category (see below). *Employment* includes “regular” employment. It does not include any form of informal or marginal employment.¹³ *Unemployment* includes periods of registered unemployment (including ALG II). The “other” category is heterogeneous. It comprises some episodes of university education, marginal employment, or other episodes in which a person was out of the labour market for other reasons.

¹² Vocational training generates “pension points”, while university education is not immediately “pension-relevant”. Although an individual does not acquire any “pension points” while in university education, these periods count as “pension-relevant periods”. An individual needs a certain minimum of pension-relevant periods before s/he can claim a pension. The maximum duration of time that can be claimed was reduced to 96 months in 1992.

¹³ Marginal employment is partially or fully exempt from social security contributions and income tax. Since 1999, the marginally employed are required to pay contributions to the public pension fund, but they are still exempt from other social security payments (such as contributions to the unemployment insurance system).

In order to capture the effect of the policy reform, we control for the *time period*. Even with large register data, the sample sizes are too small to conduct the analysis by single years, ages, and employment status and earnings. For that reason, we have grouped the calendar years into the following three broad categories: the years 2005-2006, 2007-2008, and 2009-2010. Our aim is not to measure precisely the effects of the reform; instead, we seek to map the time trends around the reform. We will return to this limitation later on.

We control for major socio-demographic characteristics. We include a binary variable that indicates whether a person was living in *eastern or western Germany*. This variable is a time-constant variable and denotes whether the person was living in eastern Germany (including East Berlin) or western Germany at the time at which they were last contacted by the German pension fund. *Age* is the baseline hazard in the event history model. It is included as a categorical variable in the model that distinguishes between ages 20-24, 25-29, 30-34, and 35-40. These cut-points have been chosen arbitrarily, but the results remain largely unchanged if more fine-grained ones are used. Moreover, other model specifications that do not make any parametric assumption regarding the baseline hazard, like the Cox model, lead to similar results.

A considerable disadvantage of this dataset is the lack of variables that could be used as control variables. While the data include information on level of education, this information is incomplete since it is provided by the employer on a voluntary basis only. However, a significant advantage of the register data is that they provide very detailed monthly *earning histories*. Earnings information is stored in terms of pension points, with one pension point constituting the average earnings in a given year. Based on this information, we have generated a variable that gives information on each individual's monthly earnings, distinguishing between low earnings (up to 50 percent of average earnings), medium earnings (50 percent to less than 100 percent of average earnings), and high earnings (100 percent or more of average earnings). This variable is a time-varying covariate that changes its value as people progress through time. As earnings generally increase with age, this variable is strongly correlated with age. We control for age, but we also provide analyses by age group.

4.2.1 Modelling strategies

The empirical investigation we conduct below relies on conventional event history modelling. In order to specify the baseline hazard, we use a piecewise constant model. Compared to the more widely used Cox model, there are several benefits to using the piecewise constant model. Most importantly, it generates estimates for the baseline hazard, which means that it provides straightforward measures for the effect of the main process time (age).¹⁴ We start with a simple model that

¹⁴ In the Cox-Model, this is only possible in a post-estimation strategy.

controls for the main socio-demographic covariates (age, year, employment, earnings, region). In the next step, we examine the effect of the policy reform by including an interaction term of the calendar year and the main covariate of interest (women's employment and earnings). The policy reform provides strong incentives to postpone childbearing until an individual's earnings are reasonably high. Thus, we assume that the association between women's employment and earnings is different before and after reform, i.e. it is assumed that stable employment and high earnings have become a prerequisite for family formation in recent years.

4.3 Results

4.3.1 *Determinants of first birth risks*

Table 1 displays the results from the event history models, with employment (Model 1) and earnings (Model 2) as the major covariates, and first birth (or rather first pregnancy) as the outcome variable. All results are provided as relative risks. The models show that birth risks were lower at younger (20-24) and older ages (35-40), but were fairly similar at ages 25-29 and 30-34. The models also indicate that the first birth rates were higher in eastern than in western Germany, reflecting the earlier family formation tendencies among eastern German women. Turning our attention to the effects of calendar time, the model suggests that the first birth rates had increased somewhat since 2005-2006. Model 1 shows that educational participation lowered the first birth rates by about 65 percent, which is in line with earlier evidence for Germany (e.g. *Andersson et al.* 2014; *Blossfeld/Huinink* 1991; *Schmitt* 2012). Unemployment (versus employment) does not seem to influence birth behaviour. It should, however, be noted that this effect varied by region, with eastern German women being more likely than their western German counterparts to postpone parenthood during unemployment (see Table A1 in the appendix).

Model 2 includes a combination factor of earnings and employment, and shows a positive gradient; i.e. that higher earnings increased the first birth transition rates. However, we cannot rule out the possibility that the strong positive effects we observe are attributable to an acceleration of birth intensities among high income earners at advanced ages. This would mean that the proportionality assumption of the model was violated, and that the effects of earnings on the birth rates varied by age. It is therefore possible that there were interaction effects of earnings and age. This aspect will be addressed in the next step of the analysis, which also focuses more narrowly on how the effects of earnings change over time.

4.3.2 *Effect heterogeneity*

The next step of the analysis contains more refined analyses by age group and calendar year. For that purpose, we have split the sample into a group of younger (ages 20-29) and older persons (ages 30-39). Furthermore, we have used the calendar year in interaction with women's earnings/employment. While an alternative strategy would be to use a three-way interaction, this would be more

Tab. 1: Results from piecewise constant model, first birth risks

	Model 1	Model 2
<i>Age</i>		
20-24	Ref.	Ref.
25-29	1.48***	1.46***
30-34	1.89***	1.84***
35-40	1.04	1.01***
<i>Region</i>		
Western Germany	Ref.	Ref.
Eastern Germany	1.18***	1.20***
<i>Years</i>		
2005-06	Ref.	Ref.
2007-08	1.05***	1.05***
2009-10	1.05***	1.05***
<i>Employment status</i>		
Education	0.34***	0.37***
Unemployment	0.96	1.08*
Other	0.60***	0.68***
Employment	Ref.	--
Earnings: low	--	Ref.
Earnings: medium	--	1.11***
Earnings: high	--	1.20***
<i>Constant (absolute risk)</i>	0.004***	0.003***
<i>Model fit</i>		
LL initial model	-15,575	-15,575
Final model	-14,473	-14,454
<i>Sample size</i>		
Persons	48,843	48,843
Events (first births)	13,913	13,913
Person-months	2,833,078	2,833,078

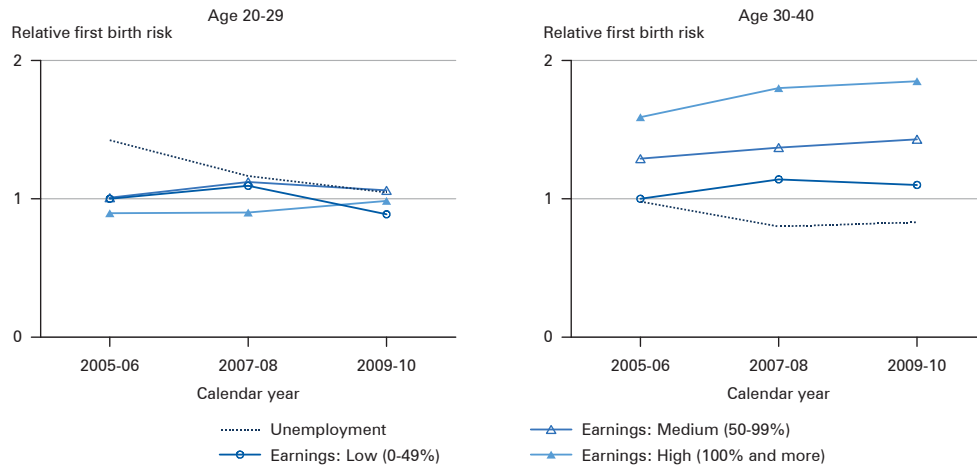
Note: Date of childbirth was backdated by nine months.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Source: VSKT 2015, own estimates

difficult to display. Figure 1 visualises the results separately for the younger and the older age groups (for the regression table, see Table A3 in the Appendix). The model corroborates earlier findings (see e.g. *Andersson et al.* 2014) showing that unemployment is positively associated with having the first birth at younger ages (below age 30), but is negatively associated with having children later in the life course. An important finding from the results is that the association between female

Fig. 1: Results from the piecewise constant model. Relative first birth risk by calendar year, age category, and female earnings



Note: Date of childbirth was backdated by nine months. Further covariates in the model are region (east/west), age (categorised and time-varying), educational participation, and other employment status in interaction with the time period.

Source: VSKT 2015, own estimates

employment, earnings, and first birth rates changed over time. After 2007, when the parental leave benefit reform came into force, the first birth rates increased among women with high earnings, while they declined or stagnated among women with lower earnings. Furthermore, the fertility pattern of unemployed women changed dramatically, with the first birth rates of younger unemployed women decreasing by 27 percent from 2005-2006 to 2009-10. While unemployment had been conducive to childbearing at younger ages before the parental leave reform came into force, it appears that this was no longer the case. We also see a change in patterns at older ages. In this age group, the first birth rates of unemployed women had dropped by 16 percent over the same time period. The results of the model provide strong evidence that the prerequisites of family formation have shifted over time in Germany. It has become increasingly important for women to have decent earnings and regular employment before starting a family.

4.3.3 Convincing evidence?

Can we attribute the pattern we observed to the parental leave benefit reform? While the findings allude to a decisive societal change, clearly attributing that shift to the parental leave regulations is more cumbersome. Causal modelling requires that “all reasonable alternative explanations (including confounders) must be ruled out” (Bhrolcháin/Dyson 2007: 9). We obviously cannot rule out the possibility that other policies – such as the increase in day care services for children that was initiated in 2005 – are the more important drivers of the changes in behavioural

patterns. An alternative explanation may also be the effects of the global financial crisis, which hit Europe at almost the same time as Germany enacted new parental leave regulations. The growing sense of economic insecurity in the wake of the crisis could have sharpened people's awareness that the single-earner model is a fragile family model – which may, in turn, have led to a reduction in first birth risks among unemployed women.

Would a more causal modelling strategy be better able to discern causal effects? A model that mapped the monthly birth rates shortly before and after the reform could provide more convincing support for the claim that there was a true “reform effect”. Indeed, these kinds of analyses have been conducted using data from the German birth registers. The results showed that there was an immediate increase in births at the cut-off date of 1 January 2007 (*Tamm* 2013). Despite the beauty of such set-ups, these findings remain limited. They may not even provide clear-cut evidence that prospective parents acted in response to the reform. Instead, they probably indicate that medical doctors and midwives have some limited leeway to influence the dates of childbirth. Thus, given that the mechanisms at the cut-off date were likely to have been very specific, these results may not be generalisable. While intricate identification strategies can easily “throw the baby out with the bathwater” (for the same analogy, see *Hill et al.* 2020: 363), the simple event history modelling approach – as illustrated above – fails to provide clear-cut causal effects, but delivers more accessible and generalisable results.

5 Conclusions

The aim of this paper was to provide an overview of recent attempts to perform causal fertility modelling based on longitudinal data. In particular, we focused on research that had tried to investigate the causal effects of female employment and family policies on birth behaviour in European countries since the 1990s. The empirical section of the paper paid special attention to the German policy reform of 2007, and raised the question of how this reform affected first birth patterns. The “reform effect” was modelled over a time-varying variable that depicted the calendar year. A main finding from this investigation was that birth rates during unemployment declined over time, whereas first birth risks increased at higher earnings levels. These results provide important insights into fertility dynamics in Germany, as they illustrate that the prerequisites of family formation have shifted rather dramatically in recent years. These shifts may be attributed to the parental leave policy reform. However, the analysis also revealed how difficult it is to rule out “alternative explanations” with simple methods, considering that the expansion of day care as well as the global financial crisis occurred around the same time as the policy was implemented, and could have caused a similar pattern.

Event history analysis is a concept that has evolved partially from demographic research, in which the aim is to portray the fertility behaviour of a given population. Event history modelling also shares the general understanding of sociological life course research (i.e. *Elder* 1985) that vital events must be situated in time

and studied along major life course dimensions (i.e. by age when first births are analysed, and by duration since the last birth when higher order births are the focus of attention). A commitment to the life course approach is often absent from many of the more causal modelling strategies. Thus, standard event history models are better able to deliver on the “life course dimension”, and to provide more accessible “descriptions of the social world” (*Brüderl/Ludwig* 2015: 353). These descriptions are of considerable value, and should probably be used as the basis before moving to the next step of analysing causal effects. But what are the reasonable next steps for causal fertility research?

The literature overview demonstrated that causal modelling approaches, and fixed-effects methods in particular, have increasingly seeped into sociological and demographic research. The wide array of European panel data, such as data from the Generations and Gender Survey, the British Understanding Society study, or the German Family Panel, have clearly fuelled this development. However, demographic and sociological scholars have been very careful when applying fixed-effects methods to fertility processes. While these methods have been regularly used to study the question of how demographic events influence other determinants, such as life satisfaction and well-being, they have rarely been used to examine birth dynamics as outcome variables. This is because in family sociology and demography, taking a parity-specific view is almost an imperative. This includes the notion that a first birth is a unique event. The analysis of first births is therefore not amenable to fixed-effects modelling since a requirement of this method is that the behaviour being studied can be repeated by the same individual under different conditions.

Other econometric methods from the programme evaluation literature, such as the regression discontinuity approach, have been employed to examine the causal effects of policy reforms on fertility behaviour. While this technique may be highly valuable in terms of advancing causal modelling in demographic research, some caution seems warranted when considering its use. It is important to keep in mind that these methods have been developed mainly for the evaluation of labour market programmes in which the outcome variables are usually frequent events, such as entering employment or receiving welfare benefits. This is clearly not the case for demographic events, which are rare. Thus, there may be only a few birth events around a cut-off date available in the survey data, which makes it difficult to generate significant results. We may intuitively prefer statistically insignificant, but well-specified results over significant and biased ones. However, an underpowered model can also raise concerns. After all, model results that are nowhere near conventional levels of significance will be difficult to interpret and hard to publish, despite all discussions and complaints about the publication bias in peer-reviewed journals. While conventional models jump to conclusions too quickly, these types of models may be characterised as being too cautious. *Hill et al.* (2020) pointed out that a “no-effect result”, which is based on a seemingly solid causal investigation, can be harmful if it leads to the discontinuation of a potentially meaningful policy measure, for instance. What seems like a trivial and purely data-driven concern is

often of considerable practical relevance, because the sample size can be a serious constraint, and because vital events are rare events.

While acknowledging that a causal perspective is important for arriving at meaningful conclusions, the paper also emphasised that the choice of method should be carefully considered when studying demographic processes. We argued that fertility research can be moved forward through more careful reflection on the origins of the bias. The observation that people act on their anticipated behaviour was regarded as a core reason why simple models that study the relationship between women's work and birth behaviour often cannot be interpreted in a causal manner. As we noted, several scholars have proposed considering the role of "narratives", which also currently seem to be gaining ground in economic research (e.g. *Bénabou et al.* 2018; *Shiller* 2020). Instead of applying ever more sophisticated econometric modelling techniques, the future of causal fertility research in this area may lay in a careful reflection of how suitable item sets could be developed to map these "imaginaries of the future self".

This overview paper has many limitations. It was restricted to fertility research based on studies that relied on longitudinal, prospective, or retrospective surveys, as well as on administrative data that were mostly collected in European countries. The paper discussed common-cause confounding and endogenous selection bias as prime sources of bias in studies that dealt with the relationship between female employment and fertility. In terms of the isolation of policy effects, several reasons were cited as to why standard tools from the programme evaluation literature cannot be used for studying fertility behaviour. Beyond the concerns listed in this paper, there are other reasons why an investigation may result in biased results. In particular, data quality may affect model results. We addressed the limitations related to small sample sizes when vital events are investigated based on survey data. However, there are numerous other aspects that revolve around the topic of "data quality". For example, recall bias in retrospective surveys can lead to biased models if events in the distant past are not remembered with the same precision as events that happened more recently. Furthermore, more salient events (childbirth) may be easier to remember than less salient ones (like events in a person's employment biography). While retrospective surveys suffer from recall bias and related problems, prospective studies suffer from attrition, which raises concerns about the selectivity of the drop-outs. Administrative data can have other problems depending on the registration system in a given country. For example, a problem of Nordic register data is that most of the information, such as information on employment status or childbirth, is only available on a yearly basis. For other countries, administrative data may not include the entire resident population. Data quality has not been addressed in this overview, but it is certainly an important additional and often overlooked aspect that warrants more attention.

Acknowledgment

For his critical feedback on an earlier version of this paper, I am thankful to Daniele Vignoli (University of Florence). I am very grateful to Felix Elwert (University of

Wisconsin) for his support in helping me to wrap my head around endogenous selection bias. Thanks also to Kai Baron (WZB) for his suggestions to engage with the economic literature on “narrative economics”. For valuable comments on matching techniques in demographic research, I would also like to thank Daniel Brüggmann (German Pension Fund). Many thanks must also go to Tatjana Mika for sharing her expertise and knowledge on the German pension data. For language editing I would like to thank Miriam Hils. All remaining errors are my own.

References

- Aassve, Arnstein et al.* 2006: Employment, family union and childbearing decisions in Great Britain. In: *Journal of the Royal Statistical Society A* 169, Part 4: 781-804. <https://doi.org/10.1111/j.1467-985X.2006.00432.x>
- Ahn, Namkee; Mira, Pedro* 2002: A note on the changing relationship between fertility and female employment rates in developed countries. In: *Journal of Political Economy* 110, 667-682. <https://doi.org/10.1007/s001480100078>
- Alderotti, Giammarco et al.* 2021: Employment instability and fertility in Europe: A meta-analysis. In: *Demography* 58,3: 871-900. <https://doi.org/10.1215/00703370-9164737>
- Allison, Paul D.* 2009: *Fixed Effects Regression Models*. Thousand Oaks: Sage.
- Andersson, Gunnar* 1999: Childbearing trends in Sweden 1961-1997. In: *European Journal of Population* 15: 1-24. <https://doi.org/10.1023/A:1006145610780>
- Andersson, Gunnar; Hoem, Jan M.; Duvander, Ann-Zofie* 2006: Social differentials in speed-premium effects in childbearing in Sweden. In: *Demographic Research* 14: 51-70. <https://doi.org/10.4054/DemRes.2006.14.4>
- Andersson, Gunnar; Kreyenfeld, Michaela; Mika, Tatjana* 2014: Welfare state context, female labour-market attachment and childbearing in Germany and Denmark. In: *Journal of Population Research* 31: 287-316. <https://doi.org/10.1007/s12546-014-9135-3>
- Annesley, Claire* 2007: Lisbon and social Europe: Towards a European 'adult worker model' welfare system. In: *Journal of European Social Policy* 17: 195-205. <https://doi.org/10.1177/0958928707078363>
- Bachrach, Christine A.; Morgan, S. Phillip* 2016: Cognitive-social model of fertility intentions. In: *Population and Development Review* 39,3: 459-485. <https://doi.org/10.1111/j.1728-4457.2013.00612.x>
- Bauernschuster, Stefan; Hener, Timo; Rainer, Helmut* 2016: Children of a (policy) revolution: The introduction of universal child care and its effect on fertility. In: *Journal of the European Economic Association* 14,4: 975-1005. <https://doi.org/10.1111/jeea.12158>
- Bazzani, Giacomo; Guetto, Raffaele; Vignoli, Daniele* 2020: Narratives of the future shape fertility in uncertain times. Evidence from the COVID-19 pandemic. DISIA Working Paper 2020/11 [https://local.disia.unifi.it/wp_disia/2020/wp_disia_2020_11.pdf, 01.07.2021].
- Beckert, Jens; Bronk, Richard* (Eds.) 2018: *Uncertain Futures: Imaginaries, Narratives, and Calculation in the Economy*. Oxford: Oxford University Press.
- Bénabou, Roland; Falk, Armin; Tirole, Jean* 2018: Narratives, imperatives, and moral Reasoning. NBER Working Paper Series 24798. <https://doi.org/10.3386/w24798>

- Bennett, Sylvana E.; Assefi, Nassim P.* 2005: School-based teenage pregnancy prevention programs: A systematic review of randomized controlled trials. In: *Journal of Adolescent Health* 36: 72-81. <https://doi.org/10.1016/j.jadohealth.2003.11.097>
- Bernardi, Laura* 2021: Qualitative longitudinal research in family sociology. In: *Schneider, Norbert F.; Kreyenfeld, Michaela* (Eds.): *Research Handbook on the Sociology of the Family*. Cheltenham: Elgar: 107-124. <https://doi.org/10.4337/9781788975544.00015>
- Berrington, Aann; Pattaro, Serena* 2014: Educational differences in fertility desires, intentions and behaviour: A life course perspective. In: *Advances in Life Course Research* 21: 10-27. <https://doi.org/10.1016/j.alcr.2013.12.003>
- Bhrolcháin, Máire Ní; Dyson, Tim* 2007: On causation in demography: Issues and illustrations. In: *Population and Development Review* 33: 1-36. <https://doi.org/10.1111/j.1728-4457.2007.00157.x>
- Björklund, Anders* 2006: Does family policy affect fertility? Lessons from Sweden. In: *Journal of Population Economics* 19: 3-24. <https://doi.org/10.1007/s00148-005-0024-0>
- Blossfeld, Hans-Peter; Huinink, Johannes* 1991: Human capital investments or norms of role transition? How women's schooling and career affect the process of family formation. In: *American Journal of Sociology* 97: 143-168. <https://doi.org/10.1086/229743>
- Boca, Daniela Del; Pasqua, Silvia; Pronzato, Chiara* 2005: Fertility and employment in Italy, France, and the UK. In: *Labour* 19: 51-77. <https://doi.org/10.1111/j.1467-9914.2005.00323.x>
- Brüderl, Josef; Ludwig, Volker* 2015: Fixed-effects panel regression. In: *Best, Hennig; Wolf, Christof* (Eds.): *The SAGE Handbook of Regression Analysis and Causal Inference*. Thousand Oaks: Sage: 327-357. <http://dx.doi.org/10.4135/9781446288146.n15>
- Castles, Francis G.* 2003: The world turned upside down: Below replacement fertility, changing preferences and family-friendly public policy in 21 OECD countries. In: *Journal of European Social Policy* 13,3: 209-227. <https://doi.org/10.1177/09589287030133001>
- Collischon, Matthias; Eberl, Andreas* 2020: Let's talk about fixed effects: Let's talk about all the good things and the bad things. In: *Kölner Zeitschrift für Soziologie und Sozialpsychologie* 72: 289-299. <https://doi.org/10.1007/s11577-020-00699-8>
- Cramer, James C.* 1980: Fertility and female employment: Problems of causal direction. In: *American Sociological Review* 45: 167-190. <https://doi.org/10.2307/2095117>
- Cygan-Rehm, Kamila* 2016: Parental leave benefit and differential fertility responses: Evidence from a German reform. In: *Journal of Population Economics* 29: 73-103. <https://doi.org/10.1007/s00148-015-0562-z>
- Cygan-Rehm, Kamila; Kuehnle, Daniel; Riphahn, Regina T.* 2018: Paid parental leave and families' living arrangements. In: *Labour Economics* 53: 182-197. <https://doi.org/10.1016/j.labeco.2018.05.008>
- Daly, Mary; Ferragina, Emanuele* 2018: Family policy in high-income countries: Five decades of development. In: *Journal of European Social Policy* 28,3: 255-270. <https://doi.org/10.1177/0958928717735060>
- Del Bono, Emilia; Weber, Andrea; Winter-Ebmer, Rudolf* 2015: Fertility and economic instability: The role of unemployment and job displacement. In: *Journal Population Economics* 28: 463-478. <https://doi.org/10.1007/s00148-014-0531-y>
- Elder Jr., Glen H.* 1985: Perspectives on the life course. In: *Elder, Glen H. Jr.* (Ed.): *Life Course Dynamics. Trajectories and Transitions 1968-1980*. Ithaca/London: Cornell University Press: 23-49.

- Elwert, Felix* 2013: Graphical causal models. In: *Morgan, Stephan* (Ed.): Handbook of Causal Analysis for Social Research. Dordrecht: Springer: 245-273. https://doi.org/10.1007/978-94-007-6094-3_13
- Elwert, Felix; Winship, Christopher* 2014: Endogenous selection bias: The problem of conditioning on a collider variable. In: *Annual Review of Sociology* 40: 31-53. <https://doi.org/10.1146/annurev-soc-071913-043455>
- Engelhardt, Henriette; Prskawetz, Alexia* 2004: On the changing correlation between fertility and female employment over space and time. In: *European Journal of Population* 20: 35-62. <https://doi.org/10.1023/B:EUJP.0000014543.95571.3b>
- Esping-Andersen, Gosta* 1999: *Social Foundations of Postindustrial Economies*. Oxford: Oxford University Press. <https://doi.org/10.1093/0198742002.001.0001>
- Farré, Lidia; González, Libertad* 2019: Does paternity leave reduce fertility? In: *Journal of Public Economics* 172: 52-56. <https://doi.org/10.1016/j.jpubeco.2018.12.002>
- Fleckenstein, Timo* 2011: The politics of ideas in welfare state transformations: Christian democracy and the reform of family policy in Germany. In: *Social Politics* 18,4: 543-571. <https://doi.org/10.1093/sp/jxr022>
- Friedman, Debra; Hechter, Michael; Kanazawa, Satoshi* 1994: A theory of the value of children. In: *Demography* 31,3: 375-401. <https://doi.org/10.2307/2061749>
- Gangl, Markus* 2010: Causal inference in sociological research. In: *Annual Review of Sociology* 36: 21-47. <https://doi.org/10.1146/annurev.soc.012809.102702>
- Gauthier, Anne H.* 2007: The impact of family policies on fertility in industrialized countries: A review of the literature. In: *Population Research Policy Review* 26: 323-346. <https://doi.org/10.1007/s11113-007-9033-x>
- Ginja, Rita; Jans, Jenny; Karimi, Arizo* 2020: Parental leave benefits, household labor supply, and children's long-run outcomes. In: *Journal of Labor Economics* 38,1. <https://doi.org/10.1086/704615>
- Hanappi, Doris et al.* 2017: Changes in employment uncertainty and the fertility intention-realization link: An analysis based on the Swiss Household Panel. In: *European Journal of Population* 33: 381-407. <https://doi.org/10.1007/s10680-016-9408-y>
- Hank, Karsten; Kreyenfeld, Michaela* 2003: A multilevel analysis of child care and women's fertility decisions in Western Germany. In: *Journal of Marriage and Family* 65,3: 584-596. <https://doi.org/10.1111/j.1741-3737.2003.00584.x>
- Hayford, Sarah R.* 2009: The evolution of fertility expectations over the life course. In: *Demography* 46,4: 765-783. <https://doi.org/10.1353/dem.0.0073>
- Hernán, Miguel A. et al.* 2002: Causal knowledge as a prerequisite for confounding evaluation: An application to birth defects epidemiology. In: *American Journal of Epidemiology* 155,2: 176-184. <https://doi.org/10.1093/aje/155.2.176>
- Hill, Terrence D. et al.* 2020: Limitations of fixed-effects models for panel data. In: *Sociological Perspectives* 63,3: 357-369. <https://doi.org/10.1177/0731121419863785>
- Hoem, Jan M.; Kreyenfeld, Michaela* 2006: Anticipatory analysis and its alternatives in life-course research. Part 1: Education and first childbearing. In: *Demographic Research* 15: 461-484. <https://dx.doi.org/10.4054/DemRes.2006.15.16>
- Hofmann, Barbara; Kreyenfeld, Michaela; Uhlendorff, Arne* 2017: Job displacement and first birth over the business cycle. In: *Demography* 54,3: 933-959. <https://doi.org/10.1007/s13524-017-0580-4>
- Huttunen, Kristiina; Kellokumpu, Jenni* 2016: The effect of job displacement on couples' fertility decisions. In: *Labour Economics* 34,2: 403-442. <https://doi.org/10.1086/683645>

- Kalwij, Adriaan* 2010: The impact of family policy expenditure on fertility in Western Europe. In: *Demography* 47,2: 503-519. <https://doi.org/10.1353/dem.0.0104>
- Krapf, Sandra* 2014: Public Childcare Provision and Fertility Behavior: A Comparison of Sweden and Germany. Opladen: Budrich University Press.
- Kravdal, Østein* 1996: How the local supply of day-care centers influences fertility in Norway: A parity-specific approach. In: *Population Research and Policy Review* 15: 201-218. <https://doi.org/10.1007/BF00127049>
- Kravdal, Østein* 2001: The high fertility of college educated women in Norway: An artefact of the separate modelling of each parity transition. In: *Demographic Research* 5: 187-215. <https://doi.org/10.4054/DemRes.2001.5.6>
- Kreyenfeld, Michaela; Mika, Tatjana* 2006: Analysemöglichkeiten der Biografiedaten des „Scientific Use Files VVL 2004“ im Bereich Fertilität und Familie. In: *Deutsche Rentenversicherung* 9-10: 583-608.
- Kuhnt, Ann-Kristin; Kreyenfeld, Michaela; Trappe, Heike* 2017: Fertility ideals of women and men across the life course. In: *Kreyenfeld, Michaela; Konietzka, Dirk* (Eds.): *Childlessness in Europe: Contexts, Causes, and Consequences*. Wiesbaden: Springer: 235-251. https://doi.org/10.1007/978-3-319-44667-7_11
- Kuhnt, Anne-Kristin; Minkus, Lara; Buhr, Petra* 2020: Uncertainty in fertility intentions from a life course perspective: Which life course markers matter? In: *Journal of Family Research* 33,1: 184-208. <https://doi.org/10.20377/jfr-426>
- Kuhnt, Ann-Kristin; Trappe, Heike* 2016: Channels of social influence on the realization of short-term fertility intentions in Germany. In: *Advances in Life Course Research* 27: 16-29. <https://doi.org/10.1016/j.alcr.2015.10.002>
- Kulu, Hill; Steele, Fiona* 2013: Interrelationships between childbearing and housing transitions in the family life course. In: *Demography* 50,5: 1687-1714. <https://doi.org/10.1007/s13524-013-0216-2>
- LaLonde, Robert J.* 1986: Evaluating the econometric evaluations of training programs with experimental data. In: *American Economic Review* 76: 604-620.
- Lee, Brian K.; Lessler, Justin; Stuart, Elizabeth A.* 2010: Improving propensity score weighting using machine learning. In: *Statistics in Medicine* 29,3: 337-346. <https://doi.org/10.1002/sim.3782>
- Lehrer, Evelyn L.; Grossbard-Shechtman, Shoshana; Leasure, J. William* 1996: Comment on 'A Theory of the Value of Children'. In: *Demography* 33,1: 133-136. <https://doi.org/10.2307/2061718>
- Lillard, Lee A.* 1993: Simultaneous equations for hazards: Marriage duration and fertility timing. In: *Journal of Econometrics* 56,1-2: 189-217. [https://doi.org/10.1016/0304-4076\(93\)90106-f](https://doi.org/10.1016/0304-4076(93)90106-f)
- Lillard, Lee A.; Panis, Constantijn W. A.* 2000: aML Multilevel Multiprocess Statistical Software, Release 1.0. Los Angeles, California: EconWare.
- Luci-Greulich, A.; Thévenon, Olivier* 2013: The impact of family policies on fertility trends in developed countries. In: *European Journal of Population* 29: 387-416. <https://doi.org/10.1007/s10680-013-9295-4>
- Matysiak, Anna; Vignoli, Daniele* 2008: Fertility and women's employment: A meta-analysis. In: *European Journal of Population* 24: 363-384. <https://doi.org/10.1007/s10680-007-9146-2>
- Matysiak, Anna; Vignoli, Daniele* 2013: Diverse effects of women's employment on fertility: Insights from Italy and Poland. In: *European Journal of Population* 29: 273-302. <https://doi.org/10.1007/s10680-013-9287-4>

- McDonald, Peter* 2000: Gender equity, social institutions and the future of fertility. In: *Journal of Population Research* 17: 1-16. <https://doi.org/10.1007/BF03029445>
- Michaud, Pierre-Carl; Tatsiramos, Konstantinos* 2011: Fertility and female employment dynamics in Europe: The effect of using alternative econometric modeling assumptions. In: *Journal of Applied Econometrics* 26,4: 641-668. <https://doi.org/10.1002/jae.1133>
- Morgan, Stephen L.; Winship, Christopher* 2015: *Counterfactuals and Causal Inference: Methods and Principles for Social Research*. 2nd edition. Cambridge: Cambridge University Press.
- Mwaikambo, Lisa et al.* 2011: What works in family planning interventions: A systematic review. In: *Studies in Family Planning* 42,2: 67-82. <https://doi.org/10.1111/j.1728-4465.2011.00267.x>
- Myrskylä, Mikko; Margolis, Rachel* 2014: Happiness: Before and after the kids. In: *Demography* 51,5: 1843-1866. <https://doi.org/10.1007/s13524-014-0321-x>
- Neyer, Gerda; Andersson, Gunnar* 2008: Consequences of family policies on childbearing behavior: Effects or artifacts? In: *Population and Development Review* 34,4: 699-724. <https://doi.org/10.1111/j.1728-4457.2008.00246.x>
- Pierson, Paul* 2000: Increasing returns, path dependence, and the study of politics. In: *American Political Science Review* 94,2: 251-267. <https://doi.org/10.2307/2586011>
- Radenacker, Anke* 2020: Changes in mothers' earnings around the time of divorce. In: *Kreyenfeld, Michaela; Trappe, Heike* (Eds.): *Parental Life Courses after Separation and Divorce in Europe*. Springer: Wiesbaden: 65-81. https://doi.org/10.1007/978-3-030-44575-1_4
- Riederer, Bernhard; Buber-Ennser, Isabella; Brzozowska, Zuzanna* 2019: Fertility intentions and their realization in couples: How the division of household chores matters. In: *Journal of Family Issues* 40,13: 1860-1882. <https://doi.org/10.1177/0192513X19848794>
- Rindfuss, Ronald R.; Brewster, Karin L.* 1996: Childrearing and fertility. In: *Population and Development Review* 22: 258-289. <https://doi.org/10.2307/2808014>
- Schmitt, Christian* 2012: A cross-national perspective on unemployment and first births. In: *European Journal of Population* 28: 303-335. <https://doi.org/10.1007/s10680-012-9262-5>
- Schneider, Daniel* 2015: The Great Recession, fertility, and uncertainty: Evidence from the United States. In: *Journal of Marriage and Family* 77,5: 1144-1156. <https://doi.org/10.1111/jomf.12212>
- Schnor, Christine* 2014: The effect of union status at first childbirth on union stability: Evidence from Eastern and Western Germany. In: *European Journal of Population* 30: 129-160. <https://doi.org/10.1007/s10680-013-9304-7>
- Schröder, Jette; Brüderl, Josef* 2008: Der Effekt der Erwerbstätigkeit von Frauen auf die Fertilität: Kausalität oder Selbstselektion? In: *Zeitschrift für Soziologie* 37,2: 117-136. <https://doi.org/10.1515/zfsoz-2008-0202>
- Sharkey, Patrick; Elwert, Felix* 2011: The legacy of disadvantage: Multigenerational neighborhood effects on cognitive ability. In: *American Journal of Sociology* 116,6: 1934-1981. <https://doi.org/10.1086/660009>
- Shiller, Robert J.* 2020: *Narrative Economics. How Stories go Viral and Drive Major Economic Events*. Princeton: Princeton University Press.
- Spéder, Zolt; Kapitány, Balázs* 2009: How are time-dependent childbearing intentions realized? Realization, postponement, abandonment, bringing forward. In: *European Journal of Population* 25: 503-523. <https://doi.org/10.1007/s10680-009-9189-7>

- Spieß, C. Katharina; Wrohlich, Katharina* 2008: The parental leave benefit reform in Germany: Costs and labour market outcomes of moving towards the Nordic model. In: *Population Research and Policy Review* 27: 575-591. <https://doi.org/10.1007/s11113-008-9086-5>
- Stegmann, Michael* 2018: Methodische Umsetzung FDZ-Biografiedatensatz – VSKT. [<https://docplayer.org/195602817-Benutzerhinweise-methodische-umsetzung-fdz-biografiedatensatz-vskt.html>, 30.06.2021].
- Stolzenberg, Ross M.; Waite, Linda J.* 1977: Age, fertility expectations and plans for employment. In: *American Sociological Review* 42,5: 769-783. <https://doi.org/10.2307/2094865>
- Tamm, Marcus* 2013: The impact of a large parental leave benefit reform on the timing of birth around the day of implementation. In: *Oxford Bulletin of Economics and Statistics* 75,4: 585-601. <https://doi.org/10.1111/j.1468-0084.2012.00707.x>
- Thelen, Kathleen* 1999: Historical institutionalism in comparative perspective. In: *Annual Review of Political Science* 2: 369-404. <https://doi.org/10.1146/annurev.polisci.2.1.369>
- Thévenon, Olivier; Gauthier, Anne H.* 2011: Family policies in developed countries: A ‘fertility-booster’ with side-effects. In: *Community, Work and Family* 14: 197-216. <https://doi.org/10.1080/13668803.2011.571400>
- Thomson, Elizabeth* 1984: Subjective utility and plans for childbearing and employment. In: *Population and Environment* 7: 198-208. <https://doi.org/10.1007/BF01255489>
- Thomson, Elizabeth* 1997: Couple childbearing desires, intentions, and births. In: *Demography* 34,3: 343-354. <https://doi.org/10.2307/3038288>
- Upchurch, Dawn M.; Lillard, Lee A.; Panis, Constantijn W. A.* 2002: Nonmarital childbearing: Influences of education, marriage, and fertility. In: *Demography* 39,2: 311-329. <https://doi.org/10.1353/dem.2002.0020>
- Vignoli, Daniele et al.* 2020a: Uncertainty and narratives of the future: A theoretical framework for contemporary fertility. In: *Schoen, Robert* (Ed.): *Analyzing Contemporary Fertility*. Cham: Springer: 25-47. https://doi.org/10.1007/978-3-030-48519-1_3
- Vignoli, Daniele et al.* 2020b: A reflection on economic uncertainty and fertility in Europe: The narrative framework. In: *Genus* 76,28. <https://doi.org/10.1186/s41118-020-00094-3>
- Waite, Linda J.; Stolzenberg, Ross M.* 1976: Intended childbearing and labor force participation of young women: Insights from nonrecursive models. In: *American Sociological Review* 41,2: 235-252. <https://doi.org/10.2307/2094471>
- Wesolowski, Katharina; Billingsley, Sunnee; Neyer, Gerda* 2020: Disentangling the complexity of family policies: SPIN data with an application to Lithuania and Sweden 1995-2015. In: *Demographic Research* 43: 1235-1262. <https://dx.doi.org/10.4054/DemRes.2020.43.42>
- Wood, Jonas; Neels, Karel* 2017: First a job, then a child? Subgroup variation in women’s employment-fertility link. In: *Advances in Life Course Research* 33, 38-52. <https://doi.org/10.1016/j.alcr.2016.09.003>

Date of submission: 14.12.2020

Date of acceptance: 31.05.2021

Prof. Dr. Michaela Kreyenfeld (✉). Hertie School. Berlin. Germany.

E-mail: Kreyenfeld@hertie-school.org

URL: <https://www.hertie-school.org/de/ueber-uns/profil/person/kreyenfeld>

Appendix

Tab. A1: Occurrence (Occ.) and exposures (Exp.) by age group

	Ages 20-29		Ages 30-39	
	Exp. (in %)	Occ.	Exp. (in %)	Occ.
<i>Age</i>				
20-24	38	1,964	--	--
25-29	62	5,573	--	--
30-34	--	--	61	4,747
35-40	--	--	39	1,629
<i>Region</i>				
Western Germany	80	5,639	84	5,371
Eastern Germany	20	1,898	16	1,005
<i>Years</i>				
2005-06	44	2,972	32	1,904
2007-08	33	2,586	33	2,176
2009-10	23	1,979	35	2,296
<i>Employment status</i>				
Education	11	290	1	45
Unemployment	7	779	7	273
Other	27	1,206	29	1,499
Earnings: low	13	1,201	9	477
Earnings: medium	30	2,998	26	1,700
Earnings: high	12	1,063	29	2,382
Person-months	1,704,821	7,537	1,128,257	6,376

Source: VSKT 2015, own estimates

Tab. A2: Results from piecewise constant model, first birth risks by eastern and western Germany

	Western Germany	Eastern Germany
<i>Age</i>		
20-24	Ref.	Ref.
25-29	1.55***	1.34***
30-34	2.10***	1.39***
35-40	1.18***	0.66***
<i>Years</i>		
1990-94	Ref.	Ref.
1995-99	1.06***	1.02
2000-06	1.07***	0.98
<i>Employment status</i>		
Education	0.33***	0.38***
Employment	Ref.	Ref.
Unemployment	1.01	0.86***
Other	0.63	0.49***
<i>Constant (absolute risk)</i>	0.004***	0.006***
<i>Model fit</i>		
LL initial model	-11,549	-3,098
Final model	-48,801	-2,864
<i>Sample size</i>		
Persons	39,689	9,154
Events (first births)	11,010	2,903
Person-months	2,309,983	523,095

Note: Date of childbirth was backdated by nine months.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Source: VSKT 2015, own estimates

Tab. A3: Results from piecewise constant model, first birth risks by age, employment status, earnings and time period

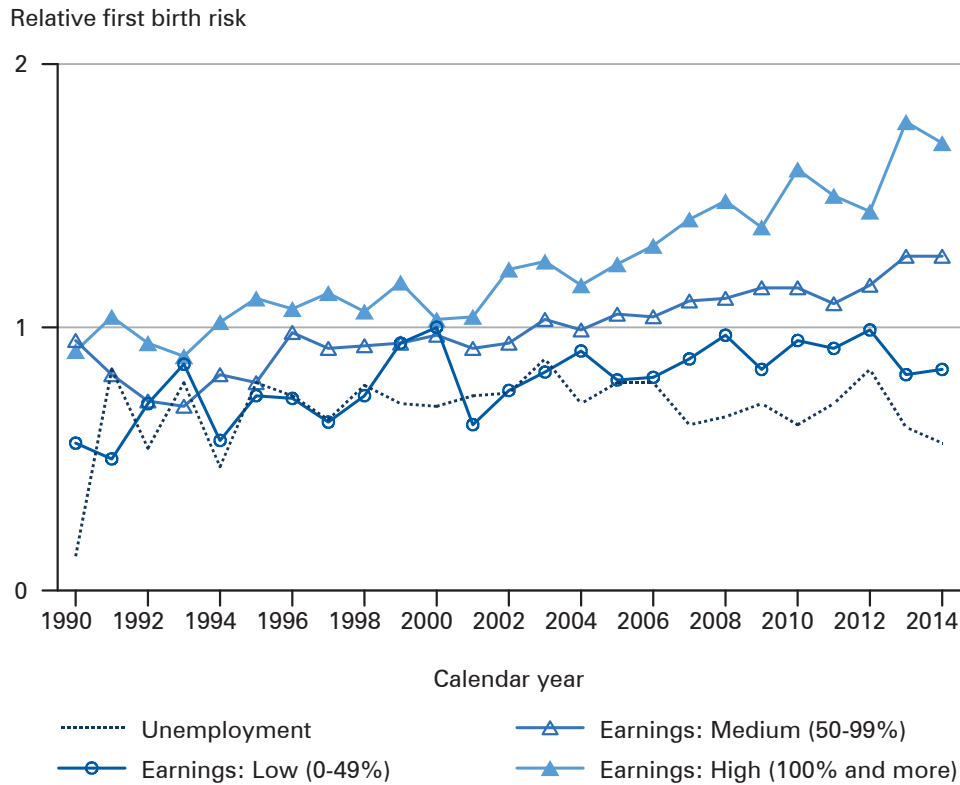
	Model 1 Age 20-29	Model 2 Age 30-40
<i>Years: 2005-06</i>		
Unemployment	1.42***	0.98
Employment: Low earnings (0-49%)	Ref.	Ref.
Employment: Medium earnings (50-99%)	1.00	1.29***
Employment: High earnings (100% and more)	0.89	1.59***
<i>Years: 2007-08</i>		
Unemployment	1.06	0.70***
Employment: Low earnings (0-49%)	Ref.	Ref.
Employment: Medium earnings (50-99%)	1.03	1.20***
Employment: High earnings (100% and more)	0.82**	1.57***
<i>Years: 2009-10</i>		
Unemployment	1.18	0.75***
Employment: Low earnings (0-49%)	Ref.	Ref.
Employment: Medium earnings (50-99%)	1.19***	1.30***
Employment: High earnings (100% and more)	1.11	1.68***

Note: Date of childbirth was backdated by nine months. Two separate models for younger ages (age 20-29) and older ages (age 30-40) were estimated. Furthermore, employment status was used in interaction with calendar year. The interaction model was re-estimated with different references categories, so that the effect of unemployment (versus low income) could more easily be compared for a given year. Further covariates in the model are region (East/West), age (categorised and time-varying) and educational participation and other employment status in interaction with the time period.

* $p < 0.05$; ** $p < 0.01$; *** $p < 0.001$

Source: VSKT 2015, own estimates

Fig. A2: Results from piecewise constant model. Relative first birth risk by single calendar year and earnings/employment status, only ages 30-40



Note: Date of childbirth was backdated by nine months. Further covariates in the model are region (East/West), age (categorised and time-varying), educational participation, and other employment status in interaction with the time period.

Source: VSKT 2015, own estimates

Comparative Population Studies

www.comparativepopulationstudies.de

ISSN: 1869-8980 (Print) – 1869-8999 (Internet)

Published by

Prof. Dr. Norbert F. Schneider

Federal Institute for Population Research
D-65180 Wiesbaden / Germany

 2021

Managing Editor

Prof. Dr. Johannes Huinink
Dr. Katrin Schiefer

Editorial Assistant

Beatriz Feiler-Fuchs
Wiebke Hamann

Layout

Beatriz Feiler-Fuchs

E-mail: cpos@bib.bund.de

Scientific Advisory Board

Karsten Hank (Cologne)
Michaela Kreyenfeld (Berlin)
Marc Luy (Vienna)
Natalie Nitsche (Rostock)
Zsolt Spéder (Budapest)
Rainer Wehrhahn (Kiel)

Board of Reviewers

Bruno Arpino (Barcelona)
Kieron Barclay (Rostock)
Laura Bernardi (Lausanne)
Gabriele Doblhammer (Rostock)
Anette Eva Fasang (Berlin)
Michael Feldhaus (Oldenburg)
Tomas Frejka (Sanibel)
Alexia Fürnkranz-Prskawetz (Vienna)
Birgit Glorius (Chemnitz)
Fanny Janssen (Groningen)
Frank Kalter (Mannheim)
Stefanie Kley (Hamburg)
Bernhard Köppen (Koblenz)
Anne-Kristin Kuhnt (Duisburg)
Hill Kulu (St Andrews)
Nadja Milewski (Wiesbaden)
Roland Rau (Rostock)
Thorsten Schneider (Leipzig)
Tomas Sobotka (Vienna)
Jeroen J. A. Spijker (Barcelona)
Heike Trappe (Rostock)
Helga de Valk (The Hague)
Sergi Vidal (Barcelona)
Michael Wagner (Cologne)